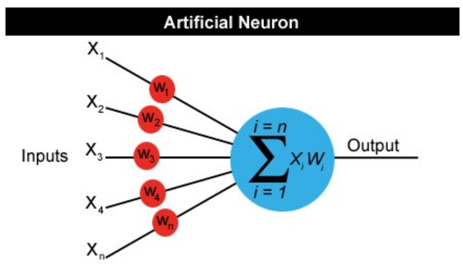
Lekcja 2: Architektura sieci neuronowych

S. Hoa Nguyen

1 Model neuronu

a) **Najważniejsze elementy:**



Rysunek 1: Model neuronu

- Wejścia + element przetwarzający + wyjście
- Wagi
- Odchylenie b (bias, offset)
- Funkcja aktywacji f

b) **Wyznaczanie sygnału wyjściowego**

- Łączny sygnał pobudzenia $net = x_1 w_1 + ... + x_n w_n + b$
- Sygnał wyjściowy $y = f(net)$

c) **Funkcje aktywacji**
Własności: Funkcje aktywacje są funkcjami niemalejącymi.

- Funkcje dyskretne:

  a) Funkcja binarna unipolarna: $f(n) = \begin{cases} 1 & \text{jeśli } n \geq 0 \\ 0 & \text{wpp.} \end{cases}$

  b) Funkcja binarna bipolarna : $f(n) = \begin{cases} 1 & \text{jeśli } n \geq 0 \\ -1 & \text{wpp.} \end{cases}$

- Funkcje ciagłe
  a) Funkcja sigmoidalna unipolarna
$$f(n) = \frac{1}{1+e^{-\lambda n}}$$
  b) Funkcja sigmoidalna bipolarna
$$f(n) = \frac{2}{1+e^{-\lambda n}} - 1$$

2 Jednowarstwowe sieci neuronowe

a) **Wyznaczanie sygnałów wyjściowych**

- Wektor sygnałów wejściowych: $\mathbf{X} = \begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{bmatrix}$
- Macierz wag:
$$\mathbf{W_{m\times n}} = \begin{bmatrix} x_{11} & x_{12} & ... & x_{1n} \\ ... & ... & ... & ... \\ x_{m1} & x_{m2} & ... & x_{mn} \end{bmatrix}$$
- Wektor odchyleń:
$$\mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ ... \\ b_m \end{bmatrix}$$
- Wektor sygnałów wyjściowych:
$$\mathbf{Y} = f(\mathbf{WX} + \mathbf{B})$$

3 Zadania podstawowe

Zadanie 1.

- Wyznaczyć sygnał wyjściowy z dwu-wejściowego neuronu zakładając, że wektor sygnałów wejściowych $X = [-2, 3]^T$, wektor wag $W = [2, -1]$, odchylenie $b = 5$, funkcja aktywacji jest dyskretna unipolarna.
- Prosta o równaniu $x_1w_1 + x_2w_2 + ...x_nw_n + b = 0(*)$ nazywa się *prostą decyzyjną*, a równianie $(*)$ *równaniem perceptronowym*. Narysować prostą decyzyjną zdefiniowaną przez podany model neuronu.

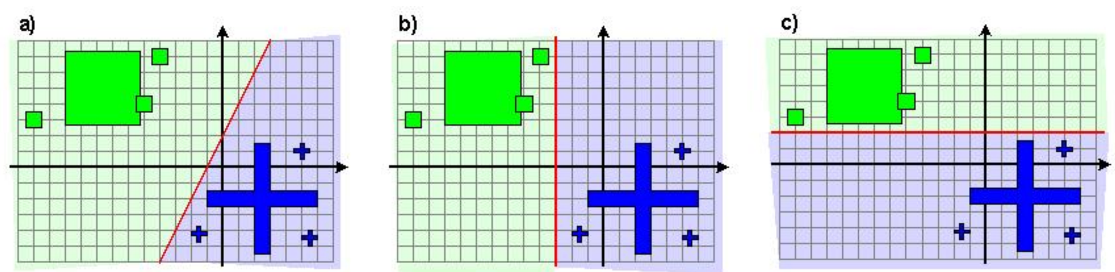
- Zakładając, że sygnały wejściowe reprezentuje współrzędne punktów na płaszczyźnie, wyznaczyć zbiór punktów, które dają sygnały wyjściowe równe 0

Zadanie 2 Zaprojektować perceptron, który oblicza następującą funkcję logiczną:

- $f(x, y) = x \wedge y$
- $f(x, y) = x \rightarrow y$
- $f(x, y) = (x \wedge y) \vee (\neg x \wedge \neg y)$.

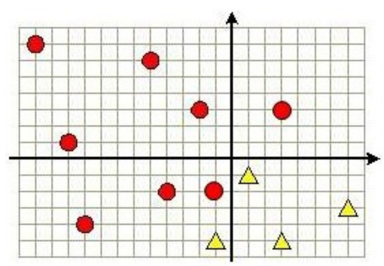
Czy perceptron może obliczyć każdą dwuargumentową funkcję logiczną?

Zadanie 3 Zbuduj trzy niezależne *dychotomizatory* (klasyfikatory dzielące zbiór danych na dwie klasy), które umożliwiają poprawną klasyfikację wszystkich punktów na płaszczyźnie zgodnie z przedstawionymi na Rysunku 2 szkicami.



Rysunek 2: Zbiór punktów do zadania 3

Zadanie 4 (1 pkt). Zbuduj sieć neuronową o dyskretnej unipolarnej funkcji aktywacji, która umożliwi poprawną klasyfikację wszystkich przedstawionych punktów podanych na Rysunku 3. Dodatkowo uwzględnij założenie, że dla punktów oznaczonych kołami na wyjściu oczekujemy wartości 1.



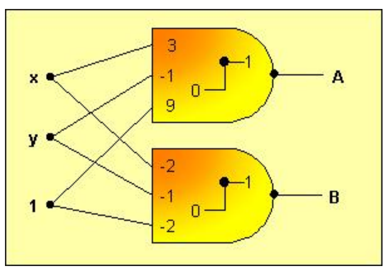
Rysunek 3: Zbiór punktów do zadania 4

Zadanie 5. Sieci neuronowej składającej z jednego neuronu użyto do klasykacji punktów w przestrzeni R^3. Neuron posiada dyskretną bipolarną funkcję aktywacji. Niech początkowy wektor wag będzie W = [-1, 2, 1], odchylenie będzie $b = -2$.

- Wyznacz sygnał wyjściowy, jeżli wektor wejściowy jest $X = (-1, 0, 3)$
- Używając reguły perceptronowej (współczynnik uczenia $\eta = 0.5$) do uczenia neuronu wyznacz nowy wektor wag po jednym cyklu uczenia, jeżeli dla wektora wejściowego $X = (-1, 0, 3)$ prawidłową odpowiedzią jest -1.
- Jaki jest błąd sieci przed i po jednym cyklu uczenia?

Zadanie 6 (1 pkt). Dla przedstawionej sieci neuronowej dla wzorca uczącego $(-1, 0)$ oczekiwanymi wartościami na wyjściach neuronów A i B są odpowiednio: 0 i 0.

- Wykonaj odpowiednie kroki algorytmu uczenia i wyznacz nowe wartości wag w neuronach.
- Jaki jest błąd sieci przed i po jednym cyklu uczenia?



Rysunek 4: Sieć neuronowa do zadania 6